DOCUMENT RESUME

ED 230 886 CG 016 753

AUTHOR Riedel, Sharon; And Others
TITLE The Response Shift Bias in Self-Report Tests: A Function of an Expectation of Change or a Shift in Internal Scaling?
PUB DATE May 82
NOTE 19p.; Paper presented at the Annual Meeting of the Midwestern Psychological Association (54th, Minneapolis, MN, May 6-8, 1982).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Assertiveness; College Students; *Counseling Effectiveness; Evaluation Criteria; Evaluation Methods; *Expectation; Females; Higher Education; *Pretests Posttests; *Response Style (Tests); *Self Evaluation (Individuals)

ABSTRACT

Self-report, pre/post testing is a frequently employed measure of therapeutic change. To investigate whether expectation of change might be an alternative explanation to the scale shift explanation of response shift bias in a self-report measure, a two-session assertiveness training intervention for college women was evaluated under manipulated conditions of expectancy or no expectancy for change. A no-treatment control group was also included. Comparison of Rathus Assertiveness Inventory pretest and retrospective pretest scores (i.e., reports of pretest status obtained at time of posttest) demonstrated a reliable response shift under the condition of expectancy for change, but not under the no expectancy for change condition. These results support a self-justification rather than scale shift explanation of response shift bias. Use of retrospective pretests to evaluate therapeutic intervention is not indicated. (Author/JAC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED230886

The Response Shift Bias in Self-Report Tests: A Function of an Expectation of Change or a Shift in Internal Scaling?

Sharon Riedel

Jana Kay Smith

Jack McKillip

Southern Illinois University at Carbondale

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Running head: Response Shift Bias

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sharon Riedel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CG 016753

Paper presented at the Annual Convention of the Midwestern Psychological Association, Minneapolis, 1982. Address comments to the first author at the Department of Psychology, Southern Illinois University, Carbondale, Illinois 62901.

The Response Shift Bias in Self-Report Tests: A Function of an Expectation of Change or a Shift in Internal Scaling?

Abstract

Alternative explanations were tested for response shift bias in self-reported measures of outcome for therapy interventions. A two session assertiveness training intervention for college women was evaluated under manipulated conditions of expectancy or no expectancy for change. A no treatment control group was also included. Comparison of Rathus Assertiveness Inventory pretest and retrospective pretest scores (i.e., reports of pretest status obtained at time of post-test) demonstrated a reliable response shift under the condition of expectancy for change, but not under the no expectancy for change condition. These results were interpreted as supporting a self-justification rather than scale shift explanation of the response shift. Use of retrospective pretests to evaluate therapeutic interventions is not indicated.

The Response Shift Bias in Self-Report Tests: A Function of an Expectation of Change or a Shift in Internal Scaling?

Veridical measurement of change attributable to a particular treatment is a central concern for program evaluators. Perhaps the most frequently employed measure of change is self-report pre/post difference scores where the difference between the posttest and pretest is taken to reflect the presence or absence of intervention effects. The popularity of pre/post measurement is partly a function of economy since the monies funded for evaluation of program effects often prohibit the incorporation of more elaborative evaluation methodologies. In addition to economy, self-report measures are often thought to provide more valid information about program effects than behavioral, objective or anecdotal measures (Howard, Schmeck & Bray, 1979).

The use of self report pre/post instruments, however, is not free of criticism. A series of studies by Howard and his colleagues (e.g., Howard, 1979; Howard, Ralph, Gulanick, Maxwell, Nance, & Gerber, 1979) reports that the self-report pretest/posttest experimental design is susceptible to an invalidity due to a change in instrumentation between the pretest and posttest. This effect they call response shift bias.

To understand the nature of the response shift bias, it is important to first understand the assumptions underlying the use of self-reported pre/post measurements. The use of self-report measures assumes that a common metric underlies both the pretest and the posttest; that is, the internal continuum which the individual refers to when making responses is the same at the posttest as it was at the pretest (Cronbach & Furby, 1970; Linn & Slinde, 1977; Lord, 1958).

However, it has been argued that a desired effect of a successful therapeutic intervention may be to disrupt that common metric or internal continuum (Cronbach & Furby, 1970; Howard, 1979; Howard, Ralph, et al., 1979; Ralph, Note 2). A treatment, by intent, may change the subjects' internal standards for judging their own behavior. The subjects' perspective or internal continuum shifts and consequently it is different at the posttest from what it was at the pretest. A pre/posttest difference score would then represent the change in the subjects' understanding of the dimensions along which they are rating themselves in addition to (or instead of) an actual change in their status. The practical conclusion from this line of reasoning is that the measuring instrument (i.e., the subject) has changed from pretest to posttest, so that differences between measurements obtained at these points in time are not readily interpretable.

The presence of a response shift bias can be identified as the difference between the retrospective pretest and the pretest (Howard, 1979) where the retrospective pretest is administered after the treatment and asks the subjects to think back and answer the questions as they would have prior to treatment. Both the pretest and retrospective pretest measure the same behavior, but if the respondent's internal scale has changed with treatment, the pretest and retrospective pretest will be using different measurement scales, and the pretest and retrospective pretest scores will consequently be different. This difference can be used then to identify response shift bias. If the response shift bias is due to a scale shift, the most appropriate measure of change or therapeutic impact would come from differences obtained using the same scale (i.e., retrospective pretest and posttest).

To better understand the response shift bias consider the following example: Before assertiveness training a subject believed she was about average in assertiveness and responded accordingly on her pretest. However, after assertiveness training the subject realized that prior to training she was really very unassertive and that now she rates herself as about average in assertiveness. Hence, pre/post differences would show that her posttest which reflects an average score does not differ from her pretest which also reflects an average score. A change score analysis would suggest that the group had no effect on the subject even though it was actually very helpful. The finding of no change would, in this case, not be due to a lack of improvement in assertiveness, but to a change in the subject's internal continuum.

Howard (1979) and Ralph (Note 2) have suggested that the use of a retrospective pretest/posttest design could overcome response shift bias when dealing with self-report instruments. After the treatment subjects are asked to respond twice to the self-report instrument -- once to describe themselves as they were before the treatment (the retrospective pretest) and once to describe themselves as they are now (the posttest). The retrospective pretest/posttest difference is then used in place of a pretest/posttest difference to measure change. Because the posttest and retrospective pretest are taken at the same point in time, the internal metric is the same for both reports and the instrumentation bias has been removed. In the above example, because the subject knew she was actually very unassertive before training, she would have rated herself lower on her retrospective pretest causing the difference score between the posttest and retrospective pretest to then accurately reflect a treatment effect.

Retrospective pretests have been shown to provide accurate measurements of treatment effect when there was no treatment effect (Howard, Daily & Gulanick, 1979; Ralph, Note 2) and when there was a treatment effect as measured by pre/post as well as retro/post differences (Dermer, Cohen, Jacobsen & Anderson, 1979; Rippey, Geller & King, 1978). Most interesting, however, is research where the retrospective pretest/posttest differences reflect a treatment effect while pre/post differences did not reflect treatment effects. Several researchers have reported the situation where the treatment did not appear beneficial when looking at pre/post differences; however, inspection of the retrospective/posttest differences did reveal treatment effects (Gulanick, Howard & Moreland, Note 1; Howard, Ralph, et al., 1979). In support of the validity of the retrospective pretest, Howard, Schmeck and Bray (1979) showed that an objective measure of change had a higher correlation with posttest/retrospective pretest differences than with posttest/pretest differences. It has been suggested that the retrospective pretest is always equal, if not superior to, the pretest in sensitivity and validity in the analysis of treatment effects (Howard, Millhan, Slaten & O'Donnell, 1981).

The above conclusions are surprising because of a long history of laboratory research in social psychology which indicates that research participants' post-experimental estimates of their pre- experimental status are inaccurate. While not mentioning the type of instrumentation artifact discussed by Howard, Ralph, et al. (1979), Nisbett and Wilson (1977) review a wide range of social psychological studies which seem to demonstrate that laboratory research subjects are completely inaccurate in the recall of their pre-experimental state. A large number of studies

are cited which demonstrate that research participants are not even aware that their emotions or attitudes have undergone dramatic changes as a result of experimental interventions. A study by Goethals and Reckman (1973) illustrates the typical laboratory findings. These investigators assessed high school students' attitudes toward busing to achieve racial integration. After the pretest, part of the group participated in a discussion designed to change their attitudes toward either pro or anti-busing. All subjects then were given a posttest and a retrospective pretest. Subjects not participating in discussion groups were able to recall their original opinions accurately. However, discussion group participants' retrospective pretest opinions shifted dramatically toward their new (posttest) opinions.

Nisbett and Wilson (1977) suggest that people have little or no direct access to complex mental processes such as those involved in judgment and decision making. They theorize that when people are asked to report their cognitive processes they do not directly access a memory of their processes but respond on the basis of some apriori causal theory. Subjects in the Goethals and Reckman study may have reasoned "if this is my opinion now, I must have had a similar opinion three weeks ago".

Research supporting the response shift bias has been conducted with volunteers participating in on-going treatment programs. These subjects had previously chosen to attend the interventions, evidencing a desire to change and an expectation that they would change. When asked to respond to the retrospective pretest, subjects may not respond on the basis of their memory or their previous state but, as Nisbett and Wilson

suggest, on the basis of an a priori theory such as "I have spent time and money on this problem, therefore, it must have had some effect. If I am average in assertiveness now, I must have been below average before the treatment." Expectation of change is an apriori causal theory of the subject and can explain the pattern of pretest, posttest and retrospective pretest differences found in previous research.

The purpose of the present study was to investigate the expectation of change as an alternative to the scale shift explanation of response shift bias in self-report measures. If the retrospective pre/pretest differences are due to a shift in the subjects' internal scaling rather than to an expectation that they should have changed as a result of the treatment, then both the no-expectation and expectation treatment groups would show a significant difference between their retrospective pretest and pretest scores. If, however, the retrospective pretest/pretest difference is significant only for an expectation treatment group, the hypothesis is supported that the difference is due to an expectation of change. In both cases, it is predicted that for a no treatment control group, the pretest will not differ from the retrospective pretest.

Method

Subjects

Subjects were female undergraduates enrolled in an introductory psychology course. They were given class credit for their participation. The treatment, assertiveness training, was given by two female graduate students in counseling psychology who had previously led workshops in assertiveness training at the university counseling center. The manipulation and all tests were administered by the two senior authors,

keeping the assertiveness trainers blind to the expectation/no expectation variable.

Self-Report Measures

The main dependent variable was assertiveness; pretest, posttest and retrospective pretests consisting of the Rathus Assertiveness Inventory (Rathus, 1973) and filler questions were administered. The two sets of items were interspersed to create one long test. The retrospective pretest differs from the pretest and posttest in that in the former subjects are asked to respond to the test items as they would have at an earlier time. Subjects' evaluation of the assertiveness training leaders was measured with the Counselor Rating Form (Barak & LaCrosse, 1975) which rated counselors on expertness, attractiveness and trustworthiness characteristics. In addition, subjects were given questions at the end of the pretest asking whether they have taken an assertiveness training course within the past six months, if they expected to change and if they expected their level of assertiveness to change as a result of the treatment.

Procedure

Subjects signed up for one of six groups, which were then randomly assigned to one of the experimental conditions. Two of the groups were no treatment control groups, two were treatment expectation of change groups and two were treatment no-expectation of change groups. The treatment consisted of two half hour assertiveness training sessions two weeks apart conducted by one of the trainers. Each trainer was randomly assigned to one of the groups within each condition. They were blind as to the experimental condition and hypotheses. The expectation of change

manipulation is described below. It was pilot tested for effectiveness and revised according to feedback provided by the pilot subjects.

In session one, the experimental groups were given the expectation or no-expectation manipulation, pretest, treatment and evaluation of the assertiveness trainer. In session two, experimental groups were given the treatment, posttest, retrospective pretest and evaluation of the assertiveness trainer. The control groups received the pretest in session one and the posttest and retrospective pretest in session two. The order of the posttest and retrospective pretests were counterbalanced within groups.

The expectation manipulation consisted of verbal and written instructions designed to create the expectation that the subject's assertive behavior would change as a result of participating in the group. The leader was presented as an expert with a great deal of experience in assertiveness training who would be using material that had great success in changing behavior. The no-expectation manipulation consisted of verbal and written instructions that did not attempt to create an expectation of change, but focused on the evaluation of the assertiveness trainer. The control groups were told that they would fill out questionnaires dealing with interpersonal communication.

Results

Manipulation Checks

Subjects' Counselor Rating Form scores for the two assertiveness training leaders were compared. Significant differences were not found between the two leaders at the end of the first or second sessions.

To determine if the two experimental groups differed in how much they expected to change as a result of the treatment, t-tests were computed between responses to the questions asking if the subjects expected to change in any way and if they expected their levels of assertiveness to change as a result of participating in the study. Significant differences were found between the expectation and the no-expectation group in their responses to both questions. The expectation groups expected both to change more and to increase their levels of assertiveness more than did the no-expectation groups ($t = 2.74$, $df = 19$, $p < .01$; and $t = 3.05$, $df = 22$, $p < .01$, respectively).

Treatment Effects

Next, the data were analyzed to determine if the treatment (i.e., assertiveness training) had an effect. Mean assertiveness scores for the three groups are presented in Table 1. Pretest/posttest difference scores were compared for the treatment (expectation and no expectation combined) and control groups. No significant differences were found, although the differences were in the predicted direction. The mean change for the treatment groups was 5.14 and -1.40 for the control groups. Similarly, when post/retrospective pretest difference scores for the two groups were compared significant differences between the treatment and control groups were not found. However, there were significant posttest-retrosepctive pretest differences for the expectation group alone (mean = 15.82; $t = 4.24$, $p < .0001$).

Table 1 about here

Response Shift

To test the main hypothesis, a series of planned comparisons were made. Pretests and retrospective pretests were compared for the expectation, no expectation and control groups. For the expectation group, the retrospective pretest scores were significantly less assertive than the pretest scores ($t = 2.94$, $df = 16$, $p < .01$). No retrospective pretest/pretest differences were found for either the no-expectation group or the control group.

Discussion

The results of this study offer support for the expectation explanation of the retrospective pretest/pretest differences. Pretest/retrospective pretest differences were found for the group that expected to change while no differences were found for either the control group or the treatment group that did not expect to change. Results for the expectation and control groups mirror the pattern Howard, Ralph et al. (1979) found in their series of studies (i.e., significant pretest/retrospective pretest differences for the treatment groups but not for the control groups). However, all of Howard's studies were performed on groups that entered treatment with the expectation that they were going to change. The failure in the present study to find pretest/retrospective pretest differences in the no-expectation group suggests that Howard's results may be attributable to expectation of change rather than response shift bias.

Analyses also revealed, as did several of Howard, Ralph et al.'s studies (1979) that while there were no significant differences between the posttest and the pretest for the treatment groups, there were signi-

ficant differences between the posttest and the retrospective pretests (for the expectation groups). Howard et al. (1979) maintain that negative findings may be explained by the response shift bias: the treatment shifted the subjects' internal measurement scale between the pre and posttest and any true differences are obscured. They maintain that the negative findings of many evaluation studies that use a pre/post self-report design can thus be explained by response shift bias and that a posttest/retrospective pretest differences scores may then be a more valid measure of treatment effects. However, results of this study show that this retrospective pretest/posttest difference occurs only in the expectation treatment group. The difference was not significant in the no-expectation group. If the retrospective pretest/posttest difference is a more valid measure of treatment effects, it should occur equally in both expectation and no-expectation treatment groups since both groups received the same treatment in the present study.

These results provide some support for the Nisbett and Wilson's (1977) observation that people are unable to directly observe or recall their cognitive processes, but infer them from a priori causal theories. Subjects who expected to change with treatment reported retrospective pretest scores significantly lower than their pretest scores, while subjects who did not expect to change showed no differences between the two sets of scores. This suggests that the subjects did not report recall prior status, rather status that was inferred from a theory of what their prior behavior should have been. However, the finding that the no-expectation and control subjects showed no differences between their pre and retrospective pretests also suggests that, in the absence

of a compelling reason to do otherwise, subjects can accurately report prior status. It may be then that people can recall prior status of mind, but if the situation provides a compelling theory about what his/her prior behavior should have been, this a priori theory overrides the actual memory.

As can be seen in Table 1, subjects in the three conditions show pretest differences. This could have been a function of the experimental procedure where the expectation manipulation was administered before the pretest. The effect of the expectation manipulation on the pretest may have differed across groups. However, it is not clear how these pretest differences could have affected the results of this study. The main hypothesis concerned pretest/retrospective pretest differences. It might be suggested that the lower expectation group pretest scores create a floor effect so that expectation subjects could not have a significantly lower retrospective pretest. However, since expectation subjects' retrospective pretests were lower than no expectation subjects' retrospective pretests, it is not likely that a floor effect accounts for the results. However, in order to correctly simulate the expectations of participants in actual treatment programs, the expectation manipulation needs to be done before pretest measurement. We are currently engaged in research examining the parameters of this pretest difference.

This study casts doubt on the scale shift explanation of differences between the retrospective pretest and pretest. Retrospective pretests cannot be recommended as a solution to response shift bias in self-report measures of treatment impact.

Reference Notes

1. Golanick, N., Howard, G. S. & Moreland, J. Evaluation of a group program for highly feminine women aimed at increased androgeny. Unpublished manuscript, University of Houston, 1977.

2. Ralph, K. Self-report response shift theory of now/then self-report procedure. Unpublished doctoral dissertation, Southern Illinois University, 1975.

References

Barak, A. & LaCrosse, M. Multidimensional perception of counselor behavior. Journal of Counseling Psychology, 1975, 22, 471-467.

Cronbach, L. & Furby, L. How we should measure change -- or should we? Psychological Bulletin, 1970, 74, 68-80.

Dermer, M., Cohen, J., Jacobsen, E. & Andersen, E. A. Evaluative judgements of aspects of life as a function of vicarious exposure to hedonic extremes. Journal of Personality and Social Psychology, 1979 37, 247-260.

Goethals, G. R. & Reckman, R. F. The perception of consistency in attitudes. Journal of Experimental Social Psychology, 1973, 9, 491-501.

Howard, G. Response shift bias: A source of contamination in self-report measures. Journal of Applied Psychology, 1979, 64, 144-150.

Howard, G., Dailey, P. & Gulanick, N. The feasibility of informated pretests in attenuating response shift bias. Applied Psychological Measurement, 1979, 3, 481-494.

Howard, G., Millham, J., Slaten, S. & O'Donnell, L. Influence of subject response style effects on retrospective measures. Applied Psychological Measurement, 1981, 5, 84-100.

Howard, G., Ralph, K., Gulanick, N., Maxwell, S., Nance, D. & Gerber, S. Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. Applied Psychological Measurement, 1979, 3, 1-23.

Table 1

Mean Assertiveness Scores for Expectation, No-Expectation, and Control Groups

Measurement	Expectation (N=17)	No-Expectation (N=12)	Control (N=10)
Pretest	2.4[a]	12.5	13.8
Retrospective Pretest	-8.4[a,b]	11.8	9.6
Post test	7.4[b]	17.7	12.4

[a]Differences between means were significant, $t=2.94$, $p < .01$

[b]Differences between means were significant, $t=4.24$, $p < .001$

Howard, G. S., Schmeck, R. R. & Bray, J. H. Internal invalidity in studies employing self-report instruments: A suggested remedy. Journal of Educational Measurement, 1979, 16, 129-135.

Linn, R. & Slinde, J. The determination of the significance of change between pre and post-testing periods. Review of Educational Research, 1977, 47, 121-150.

Lord, F. M. The measurement of growth. Educational and Psychological Measurement, 1956, 16, 421-437.

Nisbett, P. & Wilson, T. Telling more than we can know: Verbal reports on mental processes. Psychological Review, 1977, 84, 231-259.

Rathus, S. A 30 item schedule for assessing assertive behavior. Behavior Therapy, 1973, 4, 398-406.

Rippley, R., Geller, & King, D. Retrospective pretesting in the cognitive domain. Evaluation Quarterly, 1978, 2, 441-450.